Raw Sequence Listing

03/29/93 35:40:13 S543.raw

				An Aux
1			SEQUENCE LIST	ING (
2 3	(1) GEN	ERAL INFORMATIO	M·	seepp. 5,6,7,14,15,
4	(1) 021	DIGID INFORMATIO		30 pp. 0, 0, 4, 1 / 1 /
5	(i)	APPLICANT:	Donson, Jon	0-117
6			Dawson, William 0.	and 17.
7			Grantham, George L.	
8 9			Turpen, Thomas H.	* In addition, this File was not
10			Turpen, Ann Myers Garger, Stephen J.	
11			Grill, Laurence K.	tile was not
12				
13	(ii)	TITLE OF INVE	NTION: RECOMBINANT PLAN	I VIRAL NUCLEIC ACIDS SUbmiffed
14				in ASCIT.
15 16	(iii)	NUMBER OF SEQ	QUENCES: 11	11/15
17	(iv)	CORRESPONDENC	E ADDRESS:	text. Prease
18	(= - /		E: Limbach & Limbach	
19		(B) STREET:	2001 Ferry Building	ensure that
20		• • • • • • • • • • • • • • • • • • • •	n Francisco	40116
21		(D) STATE: C		los corrected
22 23		(F) ZIP: 941	.11	your corrected disk has an
24	(v)	COMPUTER READ	ABLE FORM:	cisic rus ar
25	(, ,		YPE: Floppy disk	ASCIL File
26			: IBM PC compatible	KSCIL FILE
27			G SYSTEM: PC-DOS/MS-DOS	Version #1 25
28		(D) SOFTWARE	: Patent in Release #1.	0, Version #1.25
29 30	(vi)	CURRENT APPLI	CATTON DATA .	
31	(• = /		ION NUMBER:	
32		(B) FILING D		
33		(C) CLASSIFI	CATION:	
34	(vii)	PRIOR APPLICA		
35 36			'ION NUMBER: US 600,244 PATE: 22-OCT-1990	
37		(B) FILING D	AIE: 22-0CT-1990	
38	(vii)	PRIOR APPLICA	TION DATA:	
39			ION NUMBER: US 641,617	
40		(B) FILING D	ATE: 16-JAN-1991	
41				
42 43	(vii)	PRIOR APPLICA	TION DATA:	
44			ATE: 17-FEB-1989	
45				
46	(vii)	PRIOR APPLICA		
47			TION NUMBER: US 160,766	
48		(B) FILING D	ATE: 26-FEB-1988	
49 50	(vii)	PRIOR APPLICA	TTON DATA:	
51	(+ /		TION DATA:	
52			ATE: 26-FEB-1988	

```
53
 54
     (vii)
               PRIOR APPLICATION DATA:
 55
                (A) APPLICATION NUMBER: US 347,637
 56
               (B) FILING DATE: 05-MAY-1989
 57
               PRIOR APPLICATION DATA:
 58
     (vii)
 59
                (A)
                   APPLICATION NUMBER: US 363,138
 60
               (B)
                   FILING DATE: 08-JUN-1989
 61
               PRIOR APPLICATION DATA:
     (vii)
 62
 63
               (A) APPLICATION NUMBER: US 219,279
 64
               (B) FILING DATE: 15-JUL-1988
 65
 66
               ATTORNEY/AGENT INFORMATION:
     (viii)
 67
                (A) NAME: Halluin, Albert P.
 68
                (B)
                    REGISTRATION NUMBER: 28,957
 69
               (C) REFERENCE/DOCKET NUMBER: BIOG-20121 USA
 70
 71
      (ix)
               TELECOMMUNICATION INFORMATION:
 72
                    TELEPHONE: 415-433-4150
 73
                (B)
                    TELEFAX: 415-433-8716
 74
 75
 76
     (2) INFORMATION FOR SEQ ID NO: 1:
 77
 78
       (i)
               SEQUENCE CHARACTERISTICS:
 79
                   LENGTH: 4 amino acids
 80
               (B)
                    TYPE: amino acid
 81
               (D) TOPOLOGY: linear
 82
      (ii)
 83
               MOLECULE TYPE: peptide
 84
               HYPOTHETICAL: NO
 85
     (iii)
 86
 87
      (iv)
               ANTI-SENSE: NO
 88
 89
               SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 90
 91
          Pro Xaa Gly Pro
 92
          1
 93
     (2) INFORMATION FOR SEQ ID NO: 2:
 94
 95
 96
       (i)
               SEQUENCE CHARACTERISTICS:
 97
                (A) LENGTH: 13 base pairs
 98
                (B) TYPE: nucleic acid
 99
                (C) STRANDEDNESS: single
100
                (D) TOPOLOGY: linear
101
102
      (ii)
               MOLECULE TYPE: DNA (genomic)
103
104
    (iii)
               HYPOTHETICAL: NO
```

Raw Sequence Listing

03/29/93 15:40:16 S543.raw

105			
106	(iv)	ANTI-SENSE: NO	
107			
108	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 2:	
109			
110	GGGTACC	CTGG GCC	13
111			
112			
113			
114	(2) IN	NFORMATION FOR SEQ ID NO: 3:	
115			
116	(i)	SEQUENCE CHARACTERISTICS:	
117	, – ,	(A) LENGTH: 886 base pairs	
118		(B) TYPE: nucleic acid	
119		(C) STRANDEDNESS: single	
120		(D) TOPOLOGY: linear	
121		(b) Torologi: Timear	
122	(ii)	MOLECULE TYPE: DNA (genomic)	
123	(11)	MODECOLE TIPE: DNA (genomic)	
124	(iii)	HYPOTHETICAL: NO	
125	(111)	nifoingitcal: No	
126	(iv)	ANTI-SENSE: NO	
127	(10)	ANII-SENSE: NO	
128	(vi)	ORIGINAL SOURCE:	
129	(VI)	(A) ORGANISM: Chinese cucumber	
130		(A) ORGANISM: Chinese cucumber	
131	(vii)	THORPTAME COMPONE.	
132	(\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	IMMEDIATE SOURCE:	
133		(B) CLONE: alpha-trichosanthin	
134	(ix)		
135	(IX)	FEATURE:	
136		(A) NAME/KEY: CDS (B) LOCATION: 8877	
137		(B) LOCATION: 8877	
138	(xi)	GRATIENCE DECORPRIANT CHA TO MA. 3	
139	(XI)	SEQUENCE DESCRIPTION: SEQ ID NO: 3:	
140	CTT CCT CC	7 1MG 1MG 1G1 MMG MM1 GMG GMG MGM MMG GM1 1MM GMG 1GG GMG	4.0
141	CTCGAGG	G ATG ATC AGA TTC TTA GTC CTC TCT TTG CTA ATT CTC ACC CTC	49
142		Met Ile Arg Phe Leu Val Leu Ser Leu Leu Ile Leu Thr Leu	
143		1 5 10	
144			
145	TTC CTA	A ACA ACT CCT GCT GTG GAG GGC GAT GTT AGC TTC CGT TTA TCA	97
146			
147		Thr Thr Pro Ala Val Glu Gly Asp Val Ser Phe Arg Leu Ser	
148	15	20 25 30	
149			
150	GGT GCA	A ACA AGC AGT TCC TAT GGA GTT TTC ATT TCA AAT CTG AGA AAA	145
151	~1 - T		
152	GIY Ala	A Thr Ser Ser Ser Tyr Gly Val Phe Ile Ser Asn Leu Arg Lys	
153		35 40 45	
154	00m em-		
155 156	GCT CTT	CCA AAT GAA AGG AAA CTG TAC GAT ATC CCT CTG TTA CGT TCC	193
130			

Raw Sequence Listing

03/29/93 15:40:17 S543.raw

157 158 159	Ala	Leu	Pro	Asn 50	Glu	Arg	Lys	Leu	Tyr 55	Asp	Ile	Pro	Leu	Leu 60	Arg	Ser	
160 161	TCT	CTT	CCA	GGT	TCT	CAA	CGC	TAC	GCA	TTG	ATC	CAT	CTC	ACA	AAT	TAC	241
162 163 164	Ser	Leu	Pro 65	Gly	Ser	G1n	Arg	Tyr 70	Ala	Leu	Ile	His	Leu 75	Thr	Asn	Tyr	
165 166	GCC	GAT	GAA	ACC	ATT	TCA	GTG	GCC	ATA	GAC	GTA	ACG	AAC	GTC	TAT	ATT	289
167 168 169	Ala	Asp 80	Glu	Thr	Ile	Ser	Val 85	Ala	Ile	Asp	Val	Thr 90	Asn	Val	Tyr	Ile	
170 171	ATG	GGA	TAT	CGC	GCT	GGC	GAT	ACA	TCC	TAT	TTT	TTC	AAC	GAG	GCT	TCT	337
172 173 174	Met 95	Gly	Tyr	Arg	Ala	Gly 100	Asp	Thr	Ser	Tyr	Phe 105	Phe	Asn	Glu	Ala	Ser 110	
175 176	GCA	ACA	GAA	GCT	GCA	AAA	TAT	GTA	TTC	AAA	GAC	GCT	ATG	CGA	AAA	GTT	385
177 178 179	Ala	Thr	Glu	Ala	Ala 115	Lys	Tyr	Val	Phe	Lys 120	Asp	Ala	Met	Arg	Lys 125	Val	
180 181	ACG	CTT	CCA	TAT	TCT	GGC	AAT	TAC	GAA	AGG	CTT	CAA	ACT	GCT	GCG	GGC	433
182 183 184	Thr	Leu	Pro	Tyr 130	Ser	Gly	Asn	Tyr	Glu 135	Arg	Leu	Gln	Thr	Ala 140	Ala	Gly	
185 186	AAA	ATA	AGG	GAA	AAT	ATT	CCG	CTT	GGA	CTC	CCA	GCT	TTG	GAC	AGT	GCC	481
187 188 189	Lys	Ile	Arg 145	Glu	Asn	Ile	Pro	Leu 150	Gly	Leu	Pro	Ala	Leu 155	Asp	Ser	Ala	
190 191	ATT	ACC	ACT	TTG	TTT	TAC	TAC	AAC	GCC	AAT	TCT	GCT	GCG	TCG	GCA	CTT	529
192 193 194	Ile	Thr 160	Thr	Leu	Phe	Tyr	Tyr 165	Asn	Ala	Asn	Ser	Ala 170	Ala	Ser	Ala	Leu	
195 196	ATG	GTA	CTC	ATT	CAG	TCG	ACG	TCT	GAG	GCT	GCG	AGG	TAT	AAA	TTT	ATT	577
197 198 199	M et 175	Val	Leu	Ile	Gln	Ser 180	Thr	Ser	Glu	Ala 185	Ala	Arg	Tyr	Lys	Phe	Ile 190	
200 201	GAG	CAA	CAA	ATT	GGG	AAG	CGC	GTT	GAC	AAA	ACC	TTC	CTA	CCA	AGT	TTA	625
202 203 204	G1u	Gln	Gln	Ile	Gly 195	Lys	Arg	Val	Asp	Lys 200	Thr	Phe	Leu	Pro	Ser 205	Leu	
205 206	GCA	ATT	ATA	AGT	TTG	GAA	AAT	AGT	TGG	TCT	GCT	CTC	TCC	AAG	CAA	ATT	673
207 208	Ala	Ile	Ile	Ser 210	Leu	Glu	Asn	Ser	Trp 215	Ser	Ala	Leu	Ser	Lys 220	G1n	Ile	

Raw Sequence Listing

03/29/93 15:40:19 S543.raw

209																	
210	CAG	ATA	GCG	AGT	ACT	AAT	AAT	GGA	CAG	TTT	GAA	ACT	CCT	GTT	GTG	CTT	721
211																	
212	Gln	Ile	Ala	Ser	Thr	Asn	Asn		Gln	Phe	Glu	Thr		Val	Va1	Leu	
213			225					230					235				
214																	
215	ATA	AAT	GCT	CAA	AAC	CAA	CGA	GTC	ATG	ATA	ACC	AAT	GTT	GAT	GCT	GGA	769
216 217	-1 -	3	.1.	~ 11		~ 1		7		-1	_,			_			
217	тте	240	Ala	GIN	ABN	GIN		val	Met	тте	Thr		vaı	Asp	Ala	GIÅ	
219		240					245					250					
220	CTT	CTA	ACC	TCC	220	እጥሮ	~~~	ጥጥረ	OTT C	CTC.	3 3 m	CCN	224	3 3 m	3 m/d	CC3	017
221	GII	GIA	ACC	100	AAC	AIC	GCG	116	CIG	CIG	AAI	CGA	AAC	AAT	ATG	GCA	817
222	Va1	Va 1	Thr	Ser	λan	Tla	Δla	T.011	T.011	T.011	λan	Ara	λan	λan	Mot	פוג	
223	255	• • • •		501	ADII	260	niu	cu	Deu	260	265	ALG	ADII	VOII	Mec	270	
224											203					270	
225	GCC	ATG	GAT	GAC	GAT	GTT	CCT	ATG	ACA	CAG	AGC	ттт	GGA	тст	GGA	AGT	865
226										••••						101	005
227	Ala	Met	Asp	asa	Asp	Val	Pro	Met	Thr	Gln	Ser	Phe	G1v	Cvs	Glv	Ser	
228					275					280			,	0,72	285	-0-	
229																	
230	TAT	GCT	ATT	TAG	[AAC]	CG A	\G										886
231																	
232	Tyr	Ala	Ile														
233				290)												
234																	
235																	
235 236	(2)	INI	FORM	ATION	I FOI	R SEÇ] ID	NO:	1:								
235 236 237						•	_										
235 236 237 238		INI L)	SI	guei	ICE (CHARI	- ACTEI	RIST	cs:								
235 236 237 238 239			SI (1	ZQUEI	ice (CHARI	ACTEI 289 a	RIST	cs:	lds							
235 236 237 238 239 240			SI (1	SQUEI	NCE (LENG) TYPE:	CHARI CH: 2	ACTEI 289 a	RIST: amino acid	cs:	lds							
235 236 237 238 239 240 241			SI (1	SQUEI	NCE (LENG) TYPE:	CHARI CH: 2	ACTEI 289 a	RIST: amino acid	cs:	lds							
235 236 237 238 239 240 241 242	(:	L)	SI (1 (1	EQUER A) 1 3) 7	ICE (LENGI LYPE:	CHARI TH: 2 : am: LOGY:	ACTEI 289 a ino a : 1ir	RIST: amino acid near	ICS:	ids							
235 236 237 238 239 240 241 242 243	(:		SI (1 (1	SQUEI	ICE (LENGI LYPE:	CHARI TH: 2 : am: LOGY:	ACTEI 289 a ino a : 1ir	RIST: amind acid near	ICS:	ids							
235 236 237 238 239 240 241 242 243 244	(i	i) i)	SI (1 (1 (1	EQUER A) 1 D) 1 D) 2	NCE (LENGT TYPE: TOPOI	CHARA TH: 2 : am: LOGY:	ACTEI 289 a ino a : lir	RIST: amino acid near	ICS: o ac:		JO • 4 •						
235 236 237 238 239 240 241 242 243 244 245	(:	i) i)	SI (1 (1 (1	EQUENCE COLECTION OF THE COLECTION OF TH	NCE (CENGING) TYPE: TOPOI	CHARI TH: 2 : ami LOGY: TYPE:	ACTER 289 a ino a : lir : pro	RIST: amino acid near otei:	ICS: c ac:	ID 1	10 : 4 :	:					
235 236 237 238 239 240 241 242 243 244	(i	i) i)	SI (1 (1 (1	EQUENCE COLECTION OF THE COLECTION OF TH	NCE (CENGING) TYPE: TOPOI	CHARI TH: 2 : ami LOGY: TYPE:	ACTEI 289 a ino a : lir	RIST: amino acid near otei:	ICS: c ac:	ID 1	10 : 4 :	•					
235 236 237 238 239 240 241 242 243 244 245 246 247	(i:	i) i) l)	SI (I (I MO SI	EQUER SQUER	ICE (LENGT TYPE: TOPOI TLE T NCE I	CHARACHER CHARAC	ACTER 289 a ino a : lir : pro	RIST: amind acid near otein	CCS: ac: seq	ID 1			Thr	Leu	Phe	Leu	
235 236 237 238 239 240 241 242 243 244 245 246	(i:	i) i) l)	SI (1 (1 (1	EQUER SQUER	ICE (LENGT TYPE: TOPOI TLE T NCE I	CHARACHER CHARAC	ACTER 289 a ino a : lir : pro	RIST: amind acid near otein	CCS: ac: seq	ID 1			Thr	Leu		Leu	
235 236 237 238 239 240 241 242 243 244 245 246 247 248	(ii (x:	i) i) l)	SI (I (I MO SI	EQUER SQUER	ICE (LENGT) TYPE: TOPOI	CHARACHER CHARAC	ACTER 289 a ino a : lir : pro	RIST: amind acid near otein	CCS: ac: seq	ID 1			Thr	Leu	Phe 15	Leu	
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249	(i: (x: Met	i) i) i) [][]=	SI (I (I MO SI	EQUENCE OF THE PROPERTY OF THE	ICE (CONTROL OF CONTROL OF CONTRO	CHARIANT CHEST CHE	ACTER 289 a ino a : lir : pro RIPT:	RIST: amino acid hear btein ON:	SEQ	ID 1 Leu 10	Ile	Leu			15		
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252	(i: (x: Met	i) i) i) [][]=	SI Arg	EQUENCE OF THE PROPERTY OF THE	ICE (CONTROL OF CONTROL OF CONTRO	CHARIANT CHEST CHE	ACTER 289 a ino a : lir : pro RIPT:	RIST: amino acid hear btein ON:	SEQ	ID 1 Leu 10	Ile	Leu			15		
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251	(i: (x: Met	i) i) i) [][]=	SI Arg	EQUENT SOLECTION OF THE PAIR Ala	ICE (CONTROL OF CONTROL OF CONTRO	CHARIANT CHEST CHE	ACTER 289 a ino a : lir : pro RIPT:	RIST: amino acid hear btein ON:	SEQ Leu	ID 1 Leu 10	Ile	Leu		Ser	15		
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254	(if	i) i) i) Thr	SI Arg	EQUER A) 1 3) 7 DLECT EQUER Phe	ICE (CENGRALE) TOPOI THE T	CHARI TH: 2 COGY: CYPE: DESCI Old Val	ACTER 289 a ino a : lin : pro RIPTI Leu Gly	RIST: amino acid near otein CON: Ser	SEQ Leu Val 25	ID 1 Leu 10 Ser	I1e Phe	Leu Arg	Leu	Ser 30	15 Gly	Ala	
235 236 237 238 239 240 241 242 243 244 245 246 247 248 250 251 252 253 254 255	(if	i) i) i) Thr	SI (I (I MC SI Arg	EQUER A) 1 3) 7 DLECT EQUER Phe	ICE (CENGRALE) TOPOI THE T	CHARI TH: 2 COGY: CYPE: DESCI Old Val	ACTER 289 a ino a : lin : pro RIPTI Leu Gly	RIST: amino acid near otein ON: Ser Asp	SEQ Leu Val 25	ID 1 Leu 10 Ser	I1e Phe	Leu Arg	Leu	Ser 30	15 G1y	Ala	
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256	(iii (xi Met 1 Thr	i) i) II) Thr	SI (I (I MO SI Arg Pro	EQUENCY OF THE PROPERTY OF THE	ICE (CENGRALE) TYPE: TOPOI ILE 1 Leu 5 Val	CHARIANIA CHARIA	ACTER 289 a ino a : lir : pro RIPTI Leu Gly Val	RIST: amino acid near otein CON: Ser Asp	CCS: SEQ Leu Val 25 Ile	ID 1 Leu 10 Ser	Ile Phe Asn	Leu Arg Leu	Leu Arg 45	Ser 30 Lys	15 Gly Ala	Ala Leu	
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257	(iii (xi Met 1 Thr	i) i) II) Thr Ser	SI (I (I MG SI Arg	EQUENCY OF THE PROPERTY OF THE	ICE (CENGRALE) TYPE: TOPOI ILE 1 Leu 5 Val	CHARIANIA CHARIA	ACTER 289 a ino a : lir : pro RIPTI Leu Gly Val	RIST: amino acid near otein CON: Ser Asp	CCS: SEQ Leu Val 25 Ile	ID 1 Leu 10 Ser	Ile Phe Asn	Leu Arg Leu	Leu Arg 45	Ser 30 Lys	15 Gly Ala	Ala Leu	
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258	(iii (xi Met 1 Thr	i) i) II) Thr	SI (I (I MO SI Arg Pro	EQUENCY OF THE PROPERTY OF THE	ICE (CENGRALE) TYPE: TOPOI ILE 1 Leu 5 Val	CHARIANIA CHARIA	ACTER 289 a ino a : lir : pro RIPTI Leu Gly Val	RIST: amino acid near otein CON: Ser Asp	CCS: SEQ Leu Val 25 Ile	ID 1 Leu 10 Ser	Ile Phe Asn	Leu Arg Leu	Leu Arg 45	Ser 30 Lys	15 Gly Ala	Ala Leu	
235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257	(iii (xi Met 1 Thr	i) i) Thr Ser Asn 50	SI (I (I MO SI Arg Pro	Ala 20 Ser	ICE (CENGRALE) ILE TOPOI I	CHARIANIA CHARIA	ACTER 289 a ino a : lir : pro RIPTI Leu Gly Val Tyr 55	RIST: amino acid near otein ON: Ser Asp	SEQ Leu Val 25 Ile	ID 1 Leu 10 Ser Ser	Ile Phe Asn Leu	Leu Arg Leu Leu 60	Leu Arg 45 Arg	Ser 30 Lys Ser	15 Gly Ala Ser	Ala Leu Leu	

Raw Sequence Listing

03/29/93 15:40:20 S543.raw

261	65					70					75					80	
262		_,		_													
263	Glu	Thr	Ile	Ser		Ala	Ile	Asp	Val		Asn	Va1	Tyr	Ile		G1y	
264					85					90					95		
265	_		_			_											
266	Tyr	Arg	Ala		Asp	Thr	Ser	Tyr		Phe	Asn	G1u	A1a	Ser	A1a	Thr	
267				100					105					110			
268		_	_														
269	Glu	Ala	Ala	Lys	Tyr	Val	Phe		Asp	Ala	Met	Arg	Lys	Va1	Thr	Leu	
270			115					120					125				
271	_				_	_		_	_							_	
272	Pro		Ser	Gly	Asn	Tyr		Arg	Leu	Gln	Thr		Ala	Gly	Lys	Ile	
273		130					135					140					
274	3	~1		-1.			a 1	_	_		_	_	_				
275 276		GIU	Asn	тте	Pro		GIĀ	Leu	Pro	Ala		Asp	Ser	Ala	Ile		
	145					150					155					160	
277 278	Th-	T 411	Dho	The east	Па е	3	31-	3	a			a		•			
279	Int	ьeu	Phe	Tyr		ABN	АТА	ABN	ser		ATA	ser	АТА	Leu		Val	
280					165					170					175		
281	Lou	т1 о	61 -	C	mb	C	a 1	31.	31.	3	m	T	nh.	-1 -	a 1	G1	
282	пец	116	Gln	180	1111	ser	GIU	ATA	185	Arg	Tyr	гуя	Pne	190	GIU	GIN	
283				100					103					130			
284	G1n	Tle	G1y	Tara	Ara	Va1	Aan	Tare	Thr	Dhe	T.011	D=0	go	T.OU	7.1.	T10	
285	01		195	275	A-9	V 44 1	ADD	200	1111	FIIG	пеа	PIO	205	пеа	ALG	116	
286								200					203				
287	Ile	Ser	Leu	Glu	Asn	Ser	Trp	Ser	Ala	Leu	Ser	Lvs	Gln	I1e	Gln	Tle	
288		210					215					220					
289																	
290	A1a	Ser	Thr	Asn	Asn	G1y	G1n	Phe	G1u	Thr	Pro	Va1	Va1	Leu	I1e	Asn	
291	225					230					235					240	
292																	
293	Ala	Gln	Asn	G1n	Arg	Val	Met	Ile	Thr	Asn	Val	Asp	A1a	G1y	Val	Va1	
294					245					250				_	255		
295																	
296	Thr	Ser	Asn	I1e	Ala	Leu	Leu	Leu	Asn	Arg	Asn	Asn	Met	Ala	A1a	Met	
297				260					265					270			
298																	
299	Asp	Asp	Asp	Val	Pro	Met	Thr	Gln	Ser	Phe	Gly	Cys	Gly	Ser	Tyr	Ala	
300			275					280					285				
301	_																
302	Ile																
303																	
304									_								
305	(2)	INI	ORM	ATION	1 FOI	R SE) ID	NO:	5:								
306	,,																
307	t)	L)				CHAR											
308			-			TH: (LTB	> r	-e ·	-C	hes	2 1	numberin	9
309 310			-			nuc NDEDI)
311				-		POGA :			тате								
312			,1	-, -	. 0 - 0 1			-car									

Raw Sequence Listing

03/29/93 15:40:37 S543.raw

313	(i:	i)	M	OLEC	ULE :	TYPE	: DN	A (g	enom	ic)								
314																		
315	(ii:	L)	H	YPOTI	HETI(CAL:	NO											
316																		
317	(i	7)	Al	TI-S	SENSI	E: No	0											
318																		
319	(v:	L)		RIGII														
320			(2	A) (ORGAL	NISM	: Or	yza :	sati	va								
321																		
322	(vi:	L)		MMED:														
323			(1	3) (CLONI	E: a.	lpha	amy.	lase									
324																		
325	(i)	C)		EATUI		/												
326										CATI	ON:	12.	.131	5				
327			(1	3) 1	LOCA:	LTON	: 12	1:	3 T 6									
328	1	: \				NE						_						
329	(xi	L)	SI	sQUEÌ	NCE I	UESC!	KIPT:	TOM:	SEQ	ID I	ио: !	:					_	
330	.ccm	7020	-mc -	7 200-		7 ~~	- CIT			a	a						1-	
331 332	CCT) DAD.	-1 G (_ AT(E CA	e GT(e CTC	AA e	J AC	C AT	G GT	AA ¿	C A (CAC !	TTC !	гтG	(50)	0/114
332				Mad	. dl-		1 T ^-	. >~-	, mb.	n 1/-	b 77-	1 2		. TT-1	. PL	_ T		LIS OF
334					r GII	ı va.	г тел		n Th: 5	r we	ι va.	L ASI			B Pho	e Leu	ı	الم الم
335				-	L			•	•				10	J				on this
336	TCC	Стт	ጥሮር	פידיר	כידיכי	איזיכי	GTC	СТС	Стт	ccc	CTC	TCC	TCC	אאמ	TTTC	3 C 3	98	(^ - 4 - 4 - 4 - 4 - 4 - 4 - 4 - 4 - 4 -
337			100	010		AIC	GIC	CIC	CII	GGC	CIC	100	100	AAC	110	ACA	90	line,
338	Ser	Leu	Ser	Val	Leu	Ile	Val	Leu	Leu	Glv	Leu	Ser	Ser	Asn	Len	Thr		on this. linet. Re-calwind runbering
339		15					20			1		25						ice carwing
340																	•	numbering
341	GCC	GGG	CAA	GTC	CTG	TTT	CAG	GGA	TTC	AAC	TGG	GAG	TCG	TGG	AAG	GAG	146	-)
342																	-	
343	Ala	Gly	Gln	Val	Leu	Phe	Gln	Gly	Phe	Asn	Trp	Glu	Ser	Trp	Lys	G1u		
344	30					35					40			_	_	45		
345																		
346	AAT	GGC	GGG	TGG	TAC	AAC	TTC	CTG	ATG	GGC	AAG	GTG	GAC	GAC	ATC	GCC	194	
347																		
348	Asn	Gly	Gly	Trp		Asn	Phe	Leu	Met		Lys	Val	Asp	Asp	Ile	Ala		
349					50					55					60			
350																		
351	GCA	GCC	GGC	ATC	ACC	CAC	GTC	TGG	CTC	CCT	CCG	CCG	TCT	CAC	TCT	GTC	242	
352					_,			_	_	_	_		_					
353	ALA	Ala	GTA		Thr	His	Val	Trp		Pro	Pro	Pro	Ser		Ser	Val		
354				65					70					75				
355 356	000	ar a	(12.2	000	m > ~	3 m~	aam	~~~	~~~	am-	m= ~	~-	a=-	~-	~~~	m <i>e</i> -	222	
356 357	GGC	GAG	CAA	GGC	TAC	ATG	CCT	GGG	CGG	CTG	TAC	GAT	CTG	GAC	GCG	TCT	290	
35 <i>1</i> 358	G1 • •	c1	@1 =	G1	Т	Mc+	D=-	~1. -	N	T	The	N	T	3	×1.	0		
350 359	GTÅ	GIU	80	GIA	тÅт	met	PIO	85	Arg	ьeu	Tyr	ABD	90 Leu	ABP	ATA	ser		
360			30					93					30					
361	AAG	TAC	GGC	AAC	GAG	GCG	CAG	СТС	AAG	TCG	CTG	ΔΤС	GAG	GCG	ጥጥር	СЪТ	338	
362												71.0	CAG		-10	CELL	330	
363	Lys	Tyr	Gly	Asn	G1u	Ala	Gln	Leu	Lvs	Ser	Leu	Ile	Glu	Ala	Phe	His		
364	-	95	-		_		100		4 -			105						

Raw Sequence Listing

03/29/93 15:40:44 S543.raw

365																	
366	GGC	AAG	GGC	GTC	CAG	GTG	ATC	GCC	GAC	ATC	GTC	ATC	AAC	CAC	CGC	ACG	386
367																	
368		Lys	Gly	Val	Gln	Val	Ile	Ala	Asp	Ile	Val	Ile	Asn	His	Arg	Thr	
369	110					115					120					125	
370																	
371	GCG	GAG	CAC	AAG	GAC	GGC	CGC	GGC	ATC	TAC	TGC	CTC	TTC	GAG	GGC	GGG	434
372			•	_	_				_					_			
373	Ala	Glu	His	Lys	_	Gly	Arg	Gly	Ile		Сув	Leu	Phe	Glu		Gly	
374					130					135					140		
375	100		a. a														
376	ACG	CCC	GAC	TCC	CGC	CTC	GAC	TGG	GGC	CCG	CAC	ATG	ATC	TGC	CGC	GAC	482
377 378	TT la sa	Dma		g.,	3	T		m	a 1					_	_		
379	1111	PIO	мвр		Arg	Leu	Авр	rrp		PTO	HIS	met	тте	_	Arg	Asp	
380				145					150					155			
381	GAC	ccc	TAC	ccc	CAT	GGC	N CC	ccc	220	CCC	CAC	3.00	000	000	a. a	mma	E20
382	GAC	CCC	IAC	GGC	CAI	GGC	ACC	GGC	AAC	CCG	GAC	ACC	GGC	GCC	GAC	TTC	530
383	Asn	Pro	Tur	G1 v	Agn	Gly	Thr	G1 v	λan	Dro	Acn	Thr	G1.	7 l s	N an	Pho	
384	p		160	CLY	vob	GLY	1111	165	ABII	FIO	veh	1111	170	ALG	App	FIIE	
385								103					1/0				
386	GCC	GCC	GCG	CCG	GAC	ATC	GAC	CAC	CTC	AAC	AAG	CGC	GTC	CAG	CGG	GAG	578
387																,	3.0
388	Ala	Ala	Ala	Pro	Asp	Ile	Asp	His	Leu	Asn	Lvs	Arg	Val	Gln	Arg	Glu	
389		175			-		180				4	185					
390																	
391	CTC	ATT	GGC	TGG	CTC	GAC	TGG	CTC	AAG	ATG	GAC	ATC	GGC	TTC	GAC	GCG	626
392																	
393	Leu	Ile	Gly	Trp	Leu	Asp	Trp	Leu	Lys	Met	Asp	Ile	Gly	Phe	Asp	Ala	
394	190					195					200					205	
395																	
396	TGG	CGC	CTC	GAC	TTC	GCC	AAG	GGC	TAC	TCC	GCC	GAC	ATG	GCA	AAC	ATC	674
397																	
398	Trp	Arg	Leu	Asp		Ala	Lys	Gly	Tyr		Ala	Asp	Met	Ala	_	Ile	
399					210					215					220		
400	m. a																
401 402	TAC	ATC	GAC	GCC	ACC	GAG	CCG	AGC	TTC	GCC	GTG	CCC	GAG	ATA	TCG	ACG	722
402		- 1-			m1	a 1	D		5 1						_	_,	
404	ıyr	TTE	Asp	A1a 225	Thr	Glu	PIO	ser		ата	vaı	АТА	GIU		Trp	Thr	
405				225					230					235			
406	TCC	ATG	aca	AAC	ccc	ccc	GNC	ccc	AAG	ccc	አአሮ	ጥአር	GNC	CAG	220	GCG	770
407	100	AIG	GCG	AAC	GGC	GGG	GAC	GGC	AAG	CCG	AAC	IAC	GAC	CAG	AAC	GCG	770
408	Ser	Met	λla	Δgn	Glv	Gly	Agn	Gl _V	Tare	Pro	λen	Tarr	Agn	Gln	λan	λla	
409			240		- -,	- 1		245	- , 5		*****	-1-	250	01	ADII	ALG	
410													250				
411	CAC	CGG	CAG	GAG	CTG	GTC	AAC	TGG	GTC	GAT	CGT	GTC	GGC	GGC	GCC	AAC	818
412				-	_	_	-	_	_								•
413	His	Arg	Gln	Glu	Leu	Val	Asn	Trp	Val	Asp	Arg	Val	Gly	Gly	Ala	Asn	
414		255					260	_		-	_	265	-	-			
415																	
416	ACC	AAC	GGC	ACG	GCG	TTC	GAC	TTC	ACC	ACC	AAG	GGC	ATC	CTC	AAC	GTC	866

Raw Sequence Listing

03/29/93 15:40:50 S543.raw

417																	
418	Ser	Asn	Gly	Thr	Ala	Phe	Asp	Phe	Thr	Thr	Lvs	Glv	Ile	Leu	Asn	Val	
419	270		-			275	-				280					285	
420																	
421	GCC	GTG	GAG	GGC	GAG	CTG	TGG	CGC	CTC	CGC	GGC	GAG	GAC	GGC	AAG	GCG	914
422 423	21.	7701	C1	~1	01	T	·	3	T	3	01	~1	3	a1	T -		
424	ALA	vaı	Glu	GIY	290	гел	Trp	Arg	тел	Arg 295	GIY	GIU	Авр	GIY	тув 300	Ala	
425					250					275					300		
426	CCC	GGC	ATG	ATC	GGG	TGC	TGG	CCG	GCC	AAG	GCG	ACG	ACC	TTC	GTC	GAC	962
427																	
428	Pro	Gly	Met		Gly	Trp	Trp	Pro		Lys	Ala	Thr	Thr		Val	Asp	
429 430				305					310					315			
431	AAC	CAC	GAC	ACC	GGC	TCG	ACG	CAG	CAC	СТС	тсс	CCG	ጥጥር	כככ	דרר	GAC	1010
432		0			-		1100	0.10		010	100				100	UAC	1010
433	Asn	His	Asp	Thr	Gly	Ser	Thr	Gln	His	Leu	Trp	Pro	Phe	Pro	Ser	Asp	
434			320					325					330				
435		ama		a. a			~~-										
436 437	AAG	GTC	ATG	CAG	GGC	TAC	GCA	TAC	ATC	CTC	ACC	CAC	ccc	GGC	AAC	CCA	1058
438	Lvs	Val	Met	Gln	Glv	Tvr	Ala	Tvr	Ile	Leu	Thr	His	Pro	Glv	Agn	Pro	
439	-2-	335			1	-1-	340	-1-				345		 7			
440																	
441	TGC	ATC	TTG	TAC	GAC	CAT	TTC	TTC	GAT	TGG	GGT	CTC	AAG	GAG	GAG	ATC	1106
442	~	71 4	Db -		•		5 1	n 1			~ 1	_	•	-1		_,	
443 444	350	тте	Phe	ıyr	Авр	355	Pne	Pne	Авр	Trp	360	теп	гля	GIU	GIU	365	
445	330					333					300					300	
446	GAG	CGC	CTG	GTG	TCA	ATC	AGA	AAC	CGG	CAG	GGG	ATC	CAC	CCG	GCG	AGC	1154
447																	
448	Glu	Arg	Leu	Val		Ile	Arg	Asn	Arg		Gly	Ile	His	Pro		Ser	
449 450					370					375					380		
451	GAG	CTG	CGC	ATC	ATG	GAA	GCT	GAC	AGC	GAT	CTC	TAC	СТС	GCG	GAG	ATC	1202
452																	
453	Glu	Leu	Arg		Met	Glu	Ala	Asp	Ser	Asp	Leu	Tyr	Leu	Ala	Glu	Ile	
454				385					390					395			
455 456	GAT	ggg	AAG	CTC	איזיכי	202	220	እመመ	CCA	CCN	202	ma.c	CNC	CITIC!	C 3 3	as a	1250
457	GAI	GGC	AAG	GIG	AIC	ACA	AAG	AII	GGA	CCA	AGA	IAC	GAC	GIC	GAA	CAC	1250
458	Asp	Gly	Lys	Val	Ile	Thr	Lys	Ile	Gly	Pro	Arg	Tyr	Asp	Val	Glu	His	
459	_	_	400				_	405	-		_	-	410				
460																	
461 462	CTC	ATC	CCC	GAA	GGC	TTC	CAG	GTC	GTC	GCG	CAC	GGT	GAT	GGC	TAC	GCA	1298
463	Leu	Ile	Pro	G1v	G] v	Phe	Gln	Val	Val	Ala	His	G] v	Asp	G] v	Tvr	Ala	
464		415			,		420					425		1	-1-		
465																	
466	ATC	TGG	GAG	AAA	ATC	TGA	GCGC1	ACG 1	ATGA	CGAG	AC TO	CTCA	3TTT2	A GC	AGAT:	CTAA	1353
467 468	TIA	Т∽∽	G1	Tara	LTO												
400	тте	ırp	Glu	пЛВ	тте												

Raw Sequence Listing

03/29/93 15:40:57 S543.raw

469 470															
471 472	CCTGCGATTT TTACCCTGAC CGGTATACGT ATATACGTG	C CGGCAACGAG CTGTATCCGA	1413												
473 474 475 476 477	TCCGAATTAC GGATGCAATT GTCCACGAAG TCCTCGAGG	•	1452												
478 479	· · · · · · · · · · · · · · · · · · ·														
480 481	(i) SEQUENCE CHARACTERISTICS:(A) LENGTH: 434 amino acids(B) TYPE: amino acid														
482															
483 484	(D) Topology: linear														
485 486	(ii) MOLECULE TYPE: protein														
487	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:														
488 489 490	Met Gln Val Leu Asn Thr Met Val Asn Lys Hi 1 5 10	s Phe Leu Ser Leu Ser 15													
491 492 493 494	Val Leu Ile Val Leu Leu Gly Leu Ser Ser As 20 25	n Leu Thr Ala Gly Gln 30													
495 496 497	Val Leu Phe Gln Gly Phe Asn Trp Glu Ser Tr 35 40	p Lys Glu Asn Gly Gly 45													
498 499 500	Trp Tyr Asn Phe Leu Met Gly Lys Val Asp As 50 55	p Ile Ala Ala Ala Gly 60													
501 502 503	65 70 7	s Ser Val Gly Glu Gln 5 80													
504 505 506	85 90	p Ala Ser Lys Tyr Gly 95													
507 508 509	100 105	a Phe His Gly Lys Gly 110													
510 511 512	115 120	s Arg Thr Ala Glu His 125													
513 514 515	130 135	u Gly Gly Thr Pro Asp 140													
516 517 518	145 150 15														
519 520		a Asp Phe Ala Ala Ala 175													

Raw Sequence Listing

03/29/93 15:41:04 S543.raw

521																
522	Pro	Asp	Ile	Asp	His	Leu	Asn	Lys	Arg	Val	Gln	Arg	Glu	Leu	Ile	Gly
523				180					185			_		190		-
524																
525	Trp	Leu		\mathtt{Trp}	Leu	Lys	Met	Asp	Ile	Gly	Phe	Asp	Ala	Trp	Arg	Leu
526			195					200					205			
527		_	_													
528	Asp		Ala	Lys	Gly	Tyr		Ala	Asp	Met	Ala	_	Ile	Tyr	Ile	Asp
529		210					215					220				
530				_	_											
531		Thr	Glu	Pro	ser		Ala	Val	Ala	Glu		Trp	Thr	Ser	Met	
532	225					230					235					240
533 534	3 am	~1	a 1	3	a 1	T	D	3		.	~1	•				~1
535	ABN	GIY	Gly	АВР		гув	Pro	Asn	ıyr		GIN	Asn	АТА	H18	_	GIn
536					245					250					255	
537	C1	T	7707	3	П	*** 1	3	3	**- 1	~1	a1					~ 1
538	GIU	Leu	val	260	Trp	vai	Авр	Arg	265	GIY	GIA	Ala	ABN		Asn	Gly
539				200					203					270		
540	Thr	212	Dhe	λαη	Dho	Thr	Th∽	Tazo	Glaz	Tla	Lou	λan	1751	7 T n	1701	Glu
541	1111	AIG	275	App	1116	1111	1111	280	GIY	116	пеп	ABII	285	ATA	Val	GIU
542			2,3					200					203			
543	Glv	Glu	Leu	Trp	Ara	Len	Ara	Glv	Glu	Asn	Glv	Lvs	Δla	Pro	Glv	Met
544	2	290			9		295	1			U-7	300			<u></u>	1100
545																
546	Ile	Gly	Trp	Trp	Pro	Ala	Lys	Ala	Thr	Thr	Phe	Val	asp	Asn	His	Asp
547	305	•	•	_		310	•				315					320
548																
549	Thr	Gly	Ser	Thr	Gln	His	Leu	Trp	Pro	Phe	Pro	Ser	Asp	Lys	Val	Met
550					325			_		330			_	_	335	
551																
552	Gln	Gly	Tyr	Ala	Tyr	Ile	Leu	Thr	His	Pro	Gly	Asn	Pro	Сув	Ile	Phe
553				340					345					350		
554																
555	Tyr	Asp		Phe	Phe	Asp	\mathtt{Trp}		Leu	Lys	Glu	Glu	Ile	Glu	Arg	Leu
556			355					360					365			
557					_				_			_				
558	Val		Ile	Arg	Asn	Arg		Gly	Ile	His	Pro		Ser	Glu	Leu	Arg
559		370					375					380				
560	-1-	36 - L	~ 1		•	_		_	_	_				_		_
561 562	385	Met	Glu	Ата	Авр	390	Авр	геп	ıyr	ren		GIU	тте	Asp	GIY	_
563	303					390					395					400
564	17.5	т1.	Thr	T	T1.	~1	Dwa	3	Т	3	171	a1	TT 2 _	7	-1-	D
565	val	116	1111	пув	405	GTA	PIO	AIG	ıyı	410	val	GIU	urs	ren	415	Pro
566					403					410					4 T D	
567	Glu	G) v	Phe	Gln	Va1	Va 1	Alσ	нія	Glv	Agn	Glv	ጥረታው	Δla	T10	Ттт	Glu
568				420	· 42	*41		***	425	rap	GLY	-7-	ALG	430	115	GIU
569									-27					-30		
570	Lys	Ile														
571	-															
572																

Raw Sequence Listing

03/29/93 15:41:05 S543.raw

3/3	(2) INF	ORMATION FOR SEQ ID NO:7:	
574			
575	(i)	SEQUENCE CHARACTERISTICS:	
576		(A) LENGTH: '709 base pairs	
577		(B) TYPE: nucleic acid	
578		(G) STRANDEDNESS: single	
579		(D) TOPOLOGY: linear	
580			
581	(ii)	MOLECULE TYPE: cDNA to mRNA	
582			
583	(iii)	HYPOTHETICAL: NO	
584			
585	(iv)	ANTI-SENSE: NO	
586			
587	(vi)	ORIGINAL SOURCE:	
588		(A) ORGANISM: Homo sapiens	
589			
590	(vii)	IMMEDIATE SOURCE:	
591	•	(B) CLONE: alpha-hemoglobin	
592		(, , , , , , , , , , , , , , , , , , ,	
593	(ix)	FEATURE:	
594		(A) NAME/KEY: transit peptide (B) LOCATION: 26241	
595		(B) LOCATION: 26241	
596			
597	(ix)	FEATURE:	
598		(A) NAME/KEY: CDS	
599		(B) LOCATION: 245670	
600			
601	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 7:	
602			
603	CTCGAGGG	CA TCTGATCTTT CAAGAATGGC ACAAATTAAC AACATGGCAC AAGGGATACA	60
604			
605	AACCCTTA	AT CCCAATTCCA ATTTCCATAA ACCCCAAGTT CCTAAATCTT CAAGTTTTCT	120
606			
607	TGTTTTTG	GA TGTAAAAAAC TGAAAATTC AGCAAATTCT ATGTTGGTTT TGAAAAAAGA	180
608			
609	TTCAATTT	TT ATGCAAAAGT TTTGTTCCTT TAGGATTTCA GCAGGTGGTA GAGTTTCTTG	240
610			
611	CATG GTG	CTG TCT CCT GCC GAC AAG ACC AAC GTC AAG GCC GCC TGG GGC	289
612			
613	Val	Leu Ser Pro Ala Asp Lys Thr Asn Val Lys Ala Ala Trp Cly	
614	1		
615			
616	AAG GTT (GGC GCG CAC GCT GGC GAG TAT GGT GCG GAG GCC CTG GAG AGG	337
617			
618	Lys Val	Gly Ala His Ala Gly Glu Tyr Gly Ala Glu Ala Leu Glu Arg	
619		20 25 30	
620			
621	ATG TTC	CTG TCC TTC CCC ACC ACC AAG ACC TAC TTC CCG CAC TTC GAC	385
622			
623	Met Phe 1	Leu Ser Phe Pro Thr Thr Lys Thr Tyr Phe Pro His Phe Asp	
624		35 40 45	

Raw Sequence Listing

03/29/93 15:41:17 S543.raw

625																	
626	CTG	AGC	CAC	GGC	TCT	GCC	CAG	GTT	AAG	GGC	CAC	GGC	AAG	AAG	GTG	GCC	433
627																	
628	Leu	Ser	His	Gly	Ser	Ala	Gln	Val	Lvs	Glv	His	Glv	Lvs	Lvs	Val	Ala	
629			50	•				55	•			4	60				
630																	
631	GAC	GCG	CTG	ACC	AAC	GCC	GTG	GCG	CAC	GTG	GAC	GAC	ATG	CCC	AAC	GCG	481
632																	
633	Asp .	Ala	Leu	Thr	Asn	Ala	Val	Ala	His	Val	Asp	Asp	Met	Pro	Asn	Ala	
634	_	65					70				•	75					
635																	
636	CTG	TCC	GCC	CTG	AGC	GAC	CTG	CAC	GCG	CAC	AAG	CTT	CGG	GTG	GAC	CCG	529
637																	
638	Leu	Ser	Ala	Leu	Ser	Asp	Leu	His	Ala	His	Lys	Leu	Arg	Val	Asp	Pro	
639	80					85					90		_		_	95	
640																	•
641	GTC .	AAC	TTC	AAG	CTC	CTA	AGC	CAC	TGC	CTG	CTG	GTG	ACC	CTG	GCC	GCC	577
642																	
643	Val .	Asn	Phe	Lys	Leu	Leu	Ser	His	Сув	Leu	Leu	Val	Thr	Leu	Ala	Ala	
644					100					105					110		
645																	
646	CAC	CTC	CCC	GCC	GAG	TTC	ACC	CCT	GCG	GTG	CAC	GCC	TCC	CTG	GAC	AAG	625
647																	
648	His	Leu	Pro	Ala	Glu	Phe	Thr	Pro	Ala	Val	His	Ala	Ser	Leu	Asp	Lys	
649				115					120					125			
650																	
651	TTC	CTG	GCT	TCT	GTG	AGC	ACC	GTG	CTG	ACC	TCC	AAA	TAC	CGT	TAA	GCTGGAG	677
652																	
653	Phe	Leu		Ser	Val	Ser	Thr		Leu	Thr	Ser	Lys	_	Arg			
654			130					135					140				
655													_				
656												,	i G	1	711	C101	number
657	CCTC	GGTA	LGC (CGTT	CTC	CT G	CCCG	FTCG	A CC		~	_		_ '	(O		c number
658 659													i	5	•	regu	ired at
660	(2)	TATE	ODM	. TT ()	T POI	n 079/	. TD	NO - 1					į	, \ , \		Č	7 7 2
661	(2)	TME	ORM	ATION	N FOI	K SE(ם דם	NO:	5 :				7	The		~ d	L OE THIS
662	(i	١	C1	OUE	מרים (יסגעי	A CITIEST) T @ TT	TCC.					ſ	١		
663	, -	,	_	-	LENG:		141		ino a					- 1	\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\		
664				-	CYPE:		nino			CIU	5						
665			_	-	COPOI			inear									
666			,,	, .	LOFOI	1001		.mea.	•								
667	(ii	1	м	OLECT	י א.זד	rvpr.	, · pro	htair	า								
668	,	′	***		. 200		. pr	CEIL	••								
669																	
670	(ix)	gı	ZQUEI	ICR I	ספאכיי	RIPTI	EOM•	SEO	י מד	10 · B ·	•					
671	, _,	•	51	- # - HI				- 021 •	222			•					
672	Val	Leu	Ser	Pro	Ala	Agn	Lve	Thr	Agn	Val	Lve	Alσ	Ala	Trn	Glv	Lvs	
673	1				5		_, 5			10	~10	4			15	-,, 5	
674	_				_												
675	Val	Glv	Ala	His	Ala	Glv	G1u	Tyr	Glv	Ala	Glu	Ala	Leu	Glu	Ara	Met	
676		-		20		-4			25					30		-	

Raw Sequence Listing

03/29/93 15:41:24 S543.raw

677																
678	Phe	Leu	Ser	Phe	Pro	Thr	Thr	Lys	Thr	Tyr	Phe	Pro	His	Phe	Asp	Leu
679			35					40					45			
680																
681	Ser		Gly	Ser	Ala	Gln	Val	Lys	${\tt Gly}$	His	${ t Gly}$	Lys	Lys	Val	Ala	Asp
682		50					55					60				
683	_															
684		Leu	Thr	Asn	Ala		Ala	His	Val	Asp	Asp	Met	Pro	Asn	Ala	Leu
685	65					70					75					80
686 687			•	a			•		•	_	_	_	-	_	_	
688	ser	АТА	Leu	ser	Asp 85	Leu	HIS	АТА	HIB	_	Leu	Arg	Val	Asp		
689					65					90					95	
690	Asn	Phe	Lvs	T.e.	Leu	Ser	иiя	Cva	T.e.11	T.e.11	Val	Thr	T.011	212	λla	Wie -
691			_, _	100				C, D	105	204	741		Deu	110	AIG	пть
692																
693	Leu	Pro	Ala	Glu	Phe	Thr	Pro	Ala	Val	His	Ala	Ser	Leu	Asp	Lvs	Phe
694			115					120	-				125		-1-	
695																
696	Leu	Ala	Ser	Val	Ser	Thr	Val	Leu	Thr	Ser	Lys	Tyr	Arg			
697		130					135					140				
698																
699																
700	(2)	INF	ORMA:	CION	FOR	SEQ	ID 1	10:9	:							
701	,,		۵.													
702 703	i)	L)			NCE (~		_	~	- /· -	+ page
703					ENGTI	_	_	_	•		> '	30		11-	-n .	7 7 0
705	(B) TYPE: nucleic acid															
706	(C) STRANDEDNESS: single (D) TOPOLOGY: linear															
707			`-	-, -												
708	(ii	L)	M	OLEC	ULE T	CYPE	: cDl	NA to	o mRI	A						
709																
710	(iii	L)	H	POTI	HETIC	CAL:	МО									
711																
712	(iv	7)	Al	NTI-	SENSI	3: N)									
713																
714	iv)	L)		_	NAL S											
715			(2	A) (ORGAI	NISM	: Hoi	no sa	apie	ıs						
716 717	(vii	١,	71	WED.	IATE	COTT	000.									
718	(\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \	٠,			CLONI			homo	-lob	. ~						
719			(,	٠, ،	CHOI	5. D	eca-i	Temoi	J 10D.	L11						
720	(i>	c)	F	EATU	RE:											
721	`	-,			NAME	/KEY	: tra	ansi	t per	otide	a (B) LO	CATI	ON:	26.	. 241
										,		,				
722			-		LOCA	CION	: 26	24:	L							
722 723			-			CION	: 26	24	L							
	(i)	c)	(1		LOCAT	TION	: 26	24	1							
723 724 725	(ix	c)	(1 F1 (2	B) I EATUI A) I	LOCAT RE: NAME,	/KEY	: CD:	S								
723 724 725 726	(ia	c)	(1 F1 (2	B) I EATUI A) I	LOCAT	/KEY	: CD:	S								
723 724 725	(i)		(1 F1 (1	B)] EATUI A)] B)]	LOCAT RE: NAME,	/KEY	: CD:	s 56	85							

Raw Sequence Listing

03/29/93 15:41:31 S543.raw

729				
730	CTCGAGGGGA	TCTGATCTTT CAAGAATGG	C ACAAATTAAC AACATGGCAC AAGGGATACA	60
731		·		
732	AACCCTTAAT	CCCAATTCCA ATTTCCATA	A ACCCCAAGTT CCTAAATCTT CAAGTTTTCT	120
733				
734	TGTTTTTGGA	TCTAAAAAAC TGAAAAATT	C AGCAAATTCT ATGTTGGTTT TGAAAAAAGA	180
735				
736	TTCAATTTTT	ATGCAAAAGT TTTGTTCCT	T TAGGATTTCA GCAGGTGGTA GAGTTTCTTG	240
737				
738	GATG GTG CA	.C CTG ACT CCT GAG GA	G AAG TCT GCC GTT ACT GCC CTG TGG	289
739			_	
740	_	_	u Lys Ser Ala Val Thr Ala Leu Trp	
741	1	5	10 15	
742 743	GGG >>G GTG			
744	GGC AAG GTG	AAC GIG GAI GAA GII	GGT GGT GAG GCC CTG GGC AGG CTG	337
745	Clar Larg Vol	Nam Wal Nam Clu Wal	Gly Gly Glu Ala Leu Gly Arg Leu	
746	GIY HYB VAI	20	25 30	
747		20	25 30	
748	CTG GTG GTC	TAC CCT TGG ACC CAG	AGG TTC TTT GAG TCC TTT GGG GAT	385
749			110 110 111 010 100 111 000 0M1	303
750	Leu Val Val	Tvr Pro Trp Thr Gln	Arg Phe Phe Glu Ser Phe Gly Asp	
751		35	40 45	
752				
753	CTG TCC ACT	CCT GAT GCT GTT ATG	GGC AAC CCT AAG GTG AAG GCT CAT	433
754				
755	Leu Ser Thr	Pro Asp Ala Val Met	Gly Asn Pro Lys Val Lys Ala His	
756	50	55	60	
757				
758	GGC AAG AAA	GTG CTG GGT GCC TTT	AGT GAT GGC CTG GCT CAC CTG GAC	481
759				
760 761			Ser Asp Gly Leu Ala His Leu Asp	
761 762	65	⁷⁰ Li	an extra?	
763	אאר פדר אאם		A CTG AGT GAG CTG CAC TGT GAC AAG	(₅₂₉) 5 3 6
764	AAC CIC AAG	GGC ACC 111 GCC ACC	A CIG AGI GAG CIG CAC IGI GAC AAG	323
765	Asn Leu Lvs	Gly Thr Phe Ala Thr	Leu Ser Glu Leu His Cys Asp Lys	
766	80	85	90 95	
767			,,	
768	CTG CAC GTG	GAT CCT GAG AGC TTC	AGG CTC CTA GGC AAC GTG CTG GTC	577
769				
770	Leu His Val	Asp Pro Glu Ser Phe	Arg Leu Leu Gly Asn Val Leu Val	
771		100	105 110	
772				
773	TGT GTG CTG	GCG CAT CAC TTT GGC	AAA GAA TTC ACC CCA CCA GTG CAG	625
774				
775	Cys Val Leu		Lys Glu Phe Thr Pro Pro Val Gln	
776		115	120 125	
777 770	00m 000 min		COM CMG CCM 11m CCG CTC CCC	6.70
778 779	GCT GCC TAT	CAG AAA GTG GTG GCT	GGT GTG GCT AAT GCC CTG GCC CAC	673
779 780	אום אום ייים	Cln twe Val val ala	Gly Val Ala Asn Ala Leu Ala His	
, 50	wra wra rac	GIII TAR AST AST WIR	GIY VAI AIA ABN AIA LEU AIA HIS	

Raw Sequence Listing

03/29/93 15:41:38 S543.raw

781 782		130		135	1	40	
783 784	AAG TAT	CAC TAA	GCTCGCT TTC	TTGCTGT C	CAATTTCTA TTAA	AGGTTC	722
785	Lys Tyr	His					
786 787	145	5					
788	CTTTGT	GGG TCGA	GGTCGA C				743
789							
790 791							
792	(2) IN	FORMATIC	N FOR SEQ I	D NO: 10:			
793 794	(i)	SEOIIE	NCE CHARACT	₽DT QTT ~ Q •			
795	(-)		ENGTH: 146		ds		
796			TYPE: amino				
797 798	(ii)		TOPOLOGY: 1: TULE TYPE: p:				
799			_				
800 801	(xi)	SEQUE	NCE DESCRIP	TION: SEQ	ID NO: 10:		
802	Val His	Leu Thr	Pro Glu Gl	u Lys Ser	Ala Val Thr A	la Leu Trp Gly	
803 804	1		5		10	15	
805	Lys Val	. Asn Val	. Asp Glu Va	l Gly Gly	Glu Ala Leu G	ly Arg Leu Leu	
806		20		25		30	
807 808	Val Val	. Tvr Pro	Tro Thr Gl	n Arg Phe	Phe Glu Ser P	he Gly Asp Leu	
809		35		40		45	
810 811	Ser Thr	Pro Agn	Ala Val Ma	t Gly Agn	Pro Lva Val L	ys Ala His Gly	
812	50		, Ala vai me 5	_	60	ys Ala His Gly	
813 814	T	. Wal Tan	. Glas bla ph	- 0 >	G1 T 21	• • • •	
815	65	. val Leu	70	e ser Asp	75	is Leu Asp Asn 80	
816							
817 818	Leu Lys	Gly Thr	Phe Ala Th:	r Leu Ser	Glu Leu His C	ys Asp Lys Leu 95	
819							
820 821	His Val	Asp Pro. 100		e Arg Leu 105	Leu Gly Asn V	al Leu Val Cys 110	
822							
823 824	Val Leu	Ala His	His Phe Gl	y Lys Glu 120	Phe Thr Pro P		
825		113		120	1.	25	
826					Ala Asn Ala L	eu Ala His Lys	
827 828	130	1	13	5	140		
829	Tyr His	1					
830 831	145						
832							

Raw Sequence Listing

03/29/93 15:41:44 S543.raw

833	(2)	INFORMATION FOR SEQ ID NO:11:
834		
835	(i)	SEQUENCE CHARACTERISTICS:
836		(A) LENGTH: 17 amino acids
837		(B) TYPE: amino acid
838		(D) TOPOLOGY: linear
839		
840	(ii)	MOLECULE TYPE: peptide
841		
842	(v)	FRAGMENT TYPE: N-terminal
843		. 1
844	(vi)	ORIGINAL SOURCE:
845		(A) ORGANISM: alkalophilic Bacillus sp.
846		(B) STRAIN: 38-2
847		
848	(vii)	IMMEDIATE SOURCE:
849		(B) CLONE: beta-cyclodextrin
850		12-17-1
851	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: (11) this is an I1
852		
853		Pro Asp Thr Ser Val Ser Asn Lys Gln Asn Phe Ser Thr Asp Val
854	1	5 10 15
855		
856	Ile	

PAGE: 1 SEQUENCE VERIFICATION REPORT PATENT APPLICATION US/07/923,692

DATE: 03/29/93 TIME: 15:41:46 S543

LINE ERROR

ORIGINAL TEXT

248	Wrong Amino Acid Designator	Met 11e Arg Phe Leu Val Leu Ser Le	eu Leu
245	Entered and Calc. Seq. Length differ	(x1) / SEQUENCE DESCRIPTION: SEQ	
329	Entered and Calc. Seq. Length differ	(xi) SEQUENCE DESCRIPTION: SEC	ON DI
601	Entered and Calc. Seq. Length differ	(xi) SEQUENCE DESCRIPTION: SEC	ON DI
728	Entered and Calc. Seq. Length differ	(xi) SEQUENCE DESCRIPTION: SEC	ON DI
851	Wrong Sequence Number	(xi) SEQUENCE DESCRIPTION: SE	ON DI

PAGE: 1

SEQUENCE MISSING ITEM REPORT PATENT APPLICATION US/07/923,692

DATE: 03/29/93 TIME: 15:41:46

S543

MANDATORY IDENTIFIER THAT WAS NOT FOUND

COUNTRY

PRIOR APPLICATION DATA More Identifiers Found Than Max Allowed

PAGE:

1

SEQUENCE CORRECTION REPORT PATENT APPLICATION US/07/923,692 DATE: 03/29/93 TIME: 15:41:46

S543

LINE ORIGINAL TEXT

CORRECTED TEXT

851 (xi) SEQUENCE DESCRIPTION: SEQ ID NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: IL

BACKUP/RESTORE TRANSACTION

Transaction Number:

933

03/29/93

START TIME: 15:22:19 END TIME: 15:42:01

PROCESSING TIME: 00:19:42 Input Set: S543

DOSCOPY

US/07/923,692

Translog Code Application Serial Number National PCT N Admendment

435

Application Class Application File Dt. 19920731